

Statistical Methods for Large Flight Lots and Ultra-High Reliability Applications

R. Ladbury, *Member, IEEE*, and J. L. Gorelick, *Member, IEEE*

Abstract—We present statistical techniques for evaluating random and systematic errors for use in flight performance predictions for large flight lots and ultra-high reliability applications.

Index Terms—Probability, quality assurance, radiation effects, reliability estimation, statistics.

I. INTRODUCTION

SAMPLING strategies for radiation testing often represent a compromise between generality and economy.[1] The most general strategies (those using binomial statistics) assume little about parts' radiation response distributions (RRD). However, such strategies require large samples to achieve high confidence of high success probability (e.g., 22 parts with no failures to have 90% confidence that at least 90% of the parts would pass). The main reason for these large samples is that the schemes must work for pathological thick-tailed or multimodal distributions as well as for well-behaved distributions. (The Frechet distribution in Fig. 1 is such a thick-tailed distribution.) Moreover, because binomial sampling makes no assumptions about distribution form even if we know, for example, with 90% confidence that 90% of parts pass at a dose D , we cannot say how many would pass at any other dose level. This makes it very difficult to assess the effectiveness of typical strategies for radiation hardness assurance (RHA), such as increasing radiation design margin (RDM).

More economical strategies begin with the assumption, based on technical and heuristic grounds, that radiation response within a wafer lot should be consistent from part to part [2] and assume the RRD will approximate a particular form (usually normal or lognormal). This allows the establishment of higher success probability and confidence with smaller test samples. Moreover, for small flight lots (<10 parts) and not-too-stringent reliability requirements, slight deviations of the actual RRD from the assumed form will not seriously affect RHA conclusions. However, if these conditions are violated, uncertainties in the RRD tails can dominate risk, and small samples do a poor job of constraining RRD tails or identifying pathological behavior therein.

Manuscript received July 8, 2005; revised August 26, 2005. This work was supported in part by the NASA James Webb Space Telescope (JWST) Program and in part by the NEPP Electronic Radiation Characterization (ERC) Project.

R. Ladbury is with NASA/GSFC, Greenbelt, MD 20771 USA (e-mail: Raymond.L.Ladbury@nasa.gov).

J. L. Gorelick is with Boeing Space and Intelligent Systems (BSIS), Los Angeles, CA 90009 USA (e-mail: jerry.l.gorelick@boeing.com).

Digital Object Identifier 10.1109/TNS.2005.861080

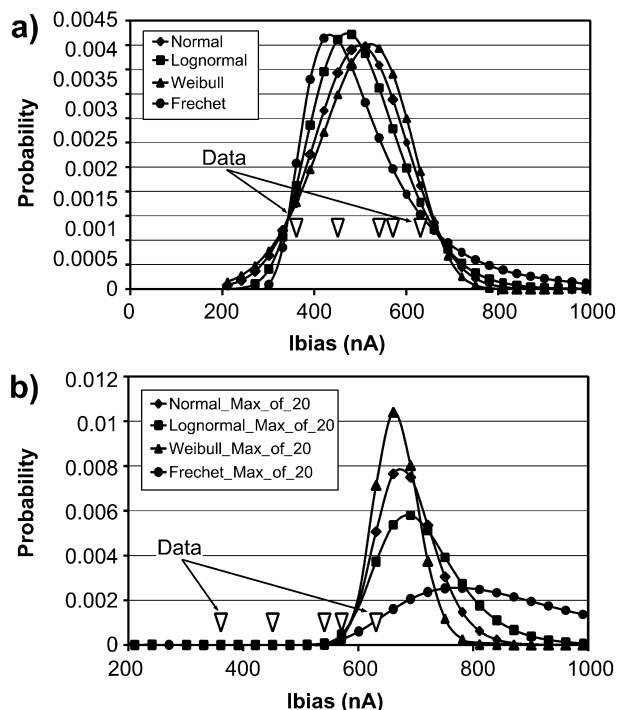


Fig. 1. (a) While normal, lognormal, Weibull, and Frechet distributions yield good fits to the data shown, they behave very differently in their tails. (b) This is clear when looking at the worst performer in flight lots as small as 20 parts.

Implications of distribution pathologies have been treated in discussions of “maverick” devices [3] and bimodality in the Enhanced Low Dose Rate Sensitivity (ELDRS) response of the National Semiconductor LM111 voltage comparator [4], [5]. Reference [3] considers the implications of occasional outliers seen in large-lot tests of 108-A op amps. Specifically, about 1% of the parts showed abnormally large changes in offset voltage. The authors of [3] concluded that such maverick devices would likely not be detected by small-sample tests, and this could have significant hardness assurance implications for some applications. Reference [4] considers the hardness assurance implications of the bimodal radiation response in National Semiconductor LM111 voltage comparators. The authors concluded that multimodality within a single wafer lot precludes sampling strategies that assume a particular distribution for any mode, since the existence of further modes at lower probability cannot be ruled out. Such parts require binomial (or, in the words of [4], “distribution-free”) sampling. The authors suggested the bimodality in the LM111 resulted from subtle differences in post-processing that shifted the balance between competing mechanisms. Reference [5] found evidence for such a competition occurring in the nitride passivation of the LM111s. The authors

concluded that the balance was shifted by differences in the preirradiation-elevated thermal stress (PETS) to which the parts were exposed after wafer fabrication. The elucidation of this mechanism meant that bimodality in the LM111 could be controlled. This reduced the urgency of developing RHA methodologies capable of dealing with pathological RRDs. However, the issue of pathological RRDs has not gone away as we show below, and, particularly for commercial parts, the remediation of such response may not be feasible. RHA methods for dealing with such parts are still needed.

In addition to such pathologies, there is also the issue of systematic errors that may be introduced when the actual RRD—though well behaved—varies from the form assumed in the analysis. As Fig. 1 shows, even well-behaved RRDs—such as the normal, Weibull, and lognormal—yield systematic errors for large flight lots under such circumstances. Similar errors occur for ultra-high-reliability applications, since here, too, uncertainties in the behavior of RRD tails can dominate risk. Usually, sample sizes in radiation characterization and radiation lot acceptance testing (RLAT) are too small to constrain RRD tails. Here, we discuss use of representative archival data to constrain distribution pathologies and provide sufficient statistics that bounding behavior for the part can be inferred. We also investigate the influence of assumed distribution form by fitting the data to several well-behaved distribution forms with different symmetries—thereby estimating the distribution dependence of the analysis.

II. DATA SOURCE

The data used in this study were mined from the BSIS radiation database and were compiled in the course of normal lot testing at the Raytheon Component Evaluation Center using their gammacell 220 Co-60 irradiator. Most parts were procured to internal BSIS Specification Control Drawing (SCD) similar to Standard Military Drawings (SMDs). All parts in the test samples were of flight quality and were burned in prior to and biased during exposure. Tests were conducted per Mil-Std. 883 Method 1019 at dose rates from 50–300 rads(Si)/s. Parametric and functional shifts were measured after each dose step. Because the test conditions and methodologies were consistent across the entire dataset, we were able to combine data for parts across many different wafer lots. Individual RLAT samples ranged from 4–20 parts.

III. METHODOLOGY

For each part, we calculated statistics at each dose step for a representative parameter for each lot and for the ensemble of all lots combined. We used the ensemble's greater statistics to characterize distribution pathologies—or if no pathologies were evident, we used binomial statistics to bound the proportion of parts that could exhibit such pathologies. Following the suggestion of Namenson [6], we varied our analysis method depending on whether variability from lot to lot greatly exceeded that typically seen within a wafer lot, or whether inter-lot and intra-lot variation were roughly commensurate. For the commensurate case, we used the ensemble to infer flight-lot behavior. When intra-lot distributions were tight, but lot-to-lot variability was large, we

combined data from the flight-lot RLAT sample and the ensemble. We used the ensemble to determine required RLAT sample sizes and likely flight-lot variability and estimated mean flight-lot performance with lot-specific tests (RLAT).

To gauge the dependence of analysis conclusions on the assumed distribution form, we fit data to three different forms. We chose the normal, Weibull, and lognormal distributions for their different symmetries, their familiarity, and the fact that they can be motivated on physical (Weibull or lognormal) or mathematical (normal) grounds:

We chose the normal distribution

$$f(x, \mu, \sigma) = (2\pi\sigma)^{-1/2} \exp(-(x - \mu)^2/\sigma^2) \quad (1)$$

because it is symmetric about its mean, it represents limiting the central behavior of most well-behaved distributions (the central limit theorem), and the sample-mean behavior for small samples is known to follow Student's *t* distribution. While it is unphysical for some problems because it is defined from $-\infty$ to $+\infty$, analysis conclusions will not be affected as long as $\mu \gg \sigma$.

We chose the Weibull distribution

$$f(x, w, s) = (s/w^s)x^{s-1} \exp(-(x/w)^s) \quad (2)$$

since for shape parameter $s > 3.68$ (corresponding to $\sigma/\mu < 0.3$), it is skewed left. For $s < 3.68$, the effects of distribution breadth will generally dwarf those of negative skew in any case. While no analog to the Student's *t* distribution exists for the Weibull, we can numerically estimate the distribution's small sample behavior as a function of s . The Weibull distribution can be motivated physically when several different failure mechanisms compete in a "race to failure."

We chose the lognormal distribution

$$f(x, \mu, \sigma) = (2\pi\sigma^2)^{-1/2} \exp(-(\ln(x) - \mu)^2/\sigma^2) \quad (3)$$

because it is skewed slightly to the right and is a common choice in radiation analyses. It can be physically motivated when damage increases in a manner proportional to damage already sustained. As with the Weibull, it is necessary to calculate numerically the small sample behavior of distribution-parameter estimators as a function of σ .

We fit data simultaneously to these three distributions using maximum likelihood (ML) techniques. The likelihood for a data set $\{x_i\}$ and a probability density function (pdf) $f(y, \theta)$ is

$$L(\{x_i\}, f(y, \theta)) = \prod_i f(x_i, \theta) \quad (4)$$

where y is the variable and θ is a scalar or vector that represents the fit parameter(s) for the pdf. The likelihood can be viewed as the probability of realizing the dataset $\{x_i\}$ if the x_i are independent and follow the distribution $f(y, \theta)$. The values of the parameter(s) θ that maximize the likelihood are called the maximum likelihood estimators (MLE) for the distribution form f and give a best fit to the data $\{x_i\}$ for that form in a maximum likelihood sense. The decrease in the relative likelihood (specifically, the logarithm of the likelihood ratio relative to the maximum likelihood) as parameter values move away from the MLEs tends to be distributed as the χ^2 distribution with degrees

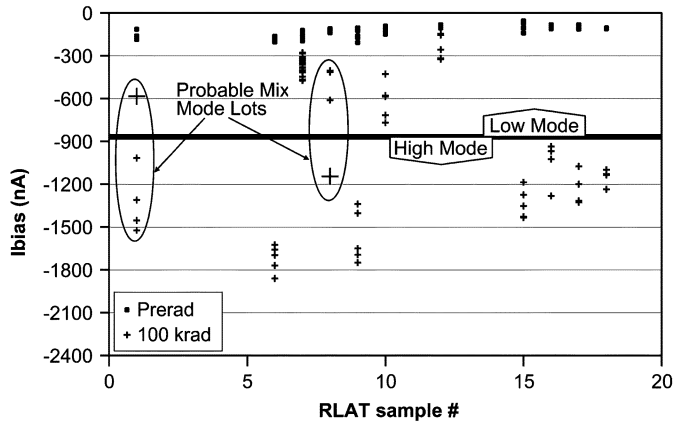


Fig. 2. OP484 Ibias versus lot and histogram after 100 krad(Si).

of freedom equal to the number of parameters for the distribution. Thus, it gives a measure of the relative probability for parameter values other than the MLEs, and one can construct confidence intervals for θ in addition to the best-fit MLEs. By taking the fit parameters that yield worst-case results and fall on the bounds of these intervals, we can bound degradation at the given confidence level.

IV. DISTRIBUTION PATHOLOGIES

Although serious pathologies in RRDs are relatively rare, we found examples of both thick-tailed and bimodal distributions. These pathologies pose problems for conventional small-sample RHA methodologies because the RRD does not approach zero in a predictable way even for very large degradation levels. This effectively invalidates strategies such as increased margin or derating, since no matter how much margin one incorporates, one cannot be confident that degradation will not exceed that level.

A. OP484

The Analog Devices OP484 quad op amp exhibits a high and a low mode of radiation-induced increased bias current, I_{bias}, with at least some lots spanning both modes (Fig. 2). While such bimodality can pose significant difficulties for RHA methodologies, the Standard Military Drawing (5962R00517) version of this part allows I_{bias} to degrade up to 3000 nA after 100 krad(Si)—much higher than the levels seen here. A study of the saturation trends in the data and of the correlations between prerad and postrad behavior suggests that, like the LM111, the bimodality here may derive from competing mechanisms and that the same mechanism responsible for the introduction of the high degradation mode may also affect the prerad leakage current.

B. 2N5019

Gate-to-source leakage current I_{gss} after 0.3 and 1 Mrad(Si) in 2N5019 junction field-effect transistors (JFETs) varies so broadly that a log scale is needed to plot it (Fig. 3). Several methods show the data may be fit by a thick-tailed (Frechet-type) extreme value or a very broad lognormal distribution (Fig. 4). Predicting on-orbit performance for such a part is difficult, since a nonzero probability exists even for very high I_{gss}.

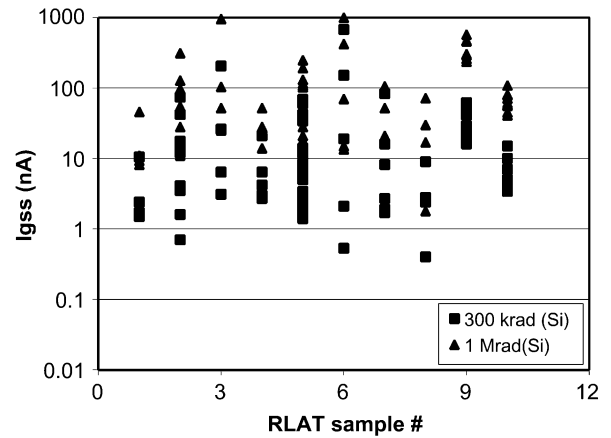


Fig. 3. I_{gss} for 2N5019 FET's spans up to 320 × in some lots.

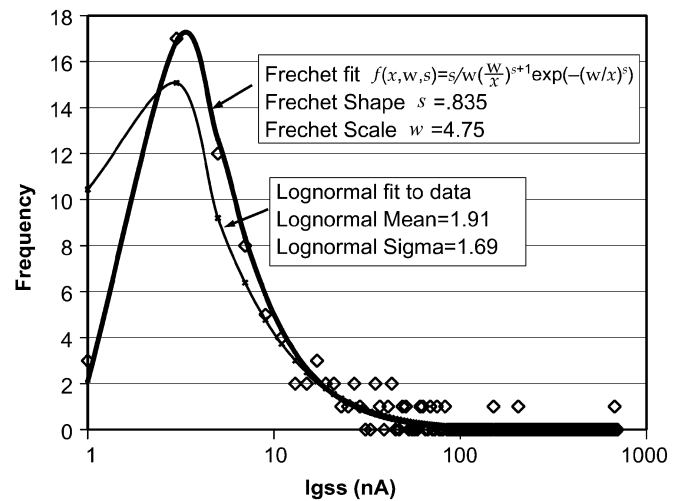


Fig. 4. I_{gss} for 2N5019 FET's fits either a thick-tailed extreme-value (Frechet-type) distribution or a very broad lognormal.

C. 2N2658

The radiation-induced gain change in the 2N2658 bipolar junction transistor (BJT) (Fig. 5) is also thick-tailed, following either a broad lognormal or Frechet distribution. If individual lots of the 2N2658 had been considered in isolation, several parts falling in the tail of the distribution could have been dismissed as “outliers.” Instead, the improved statistics of the ensemble reveal that they are part of distribution pathologies rather than merely isolated occurrences. As such, it is inherently risky to try to establish derating guidelines—let alone decide whether to waive RLAT—based on results for one or a few lots.

For parts that exhibit RRD pathologies, predicting the likely amount of degradation is inherently difficult. Therefore, the use of such parts cannot be recommended unless the application is insensitive to the types of degradation observed.

V. WELL-BEHAVED DATA

A. RH1014 Op amp

To illustrate the use of well-behaved archival data, we use 38 RLAT samples (158 parts total) of the Linear Technologies RH1014 quad op amp tested at 60, 100, and 200 krad(Si). The

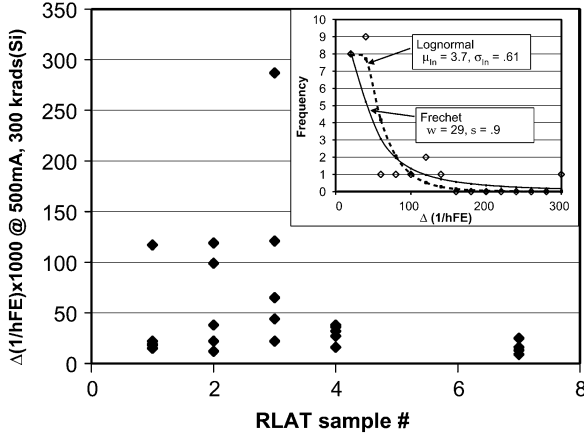


Fig. 5. Changes in gain of 2N2658 BJTs after 300 krad(Si) span more than $32 \times$. The inset shows the histogram of gain change, along with Lognormal and Frechet fits to the data.

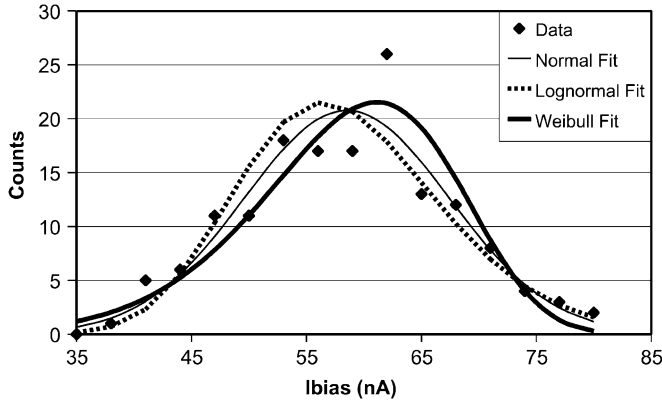


Fig. 6. Ibias after 200 krad(Si) and best fits to the data.

fact that the 158-part dataset yielded no sign of RRD pathology establishes at the 90% CL (using binomial statistics) that any such pathologies comprise less than 1.5% of the parts. Fig. 6 shows Ibias after 200 krad(Si), along with normal, lognormal, and Weibull fits to the data.

ML fits to the 200-krad data yield confidence intervals (the shaded ellipses in Fig. 7) as well as best-fit values (black). For a desired confidence level CL, we select the fit parameters that yield the worst-case (WC) distribution and fall on or within the CL contour. This defines the WC distribution consistent with our data at confidence level CL. For desired success probability P_s , we define the design-to Ibias for our WC distribution as the current where we have probability P_s that we will not exceed that value. As shown in Table I, the required design-to Ibias increases with both CL and P_s .

Contributions to this increase in design-to Ibias include both random errors due to sampling and systematic errors arising from inconsistencies between the data and the form assumed by the fit. While we are usually interested in bounding degradation, in some cases, it may be desirable to separate the likely contributions of random and systematic errors. For example, we may wish to assess whether our model yields unacceptably high systematic errors or to determine test sample sizes. We first estimate the random sampling errors for an RRD of the assumed

form (e.g., normal, Weibull, or lognormal) with our ML fit parameters. We assume that the residual increase arises from systematic errors.

For the mean of a normal distribution, the sampling errors are well defined in terms of the student's t distribution with a standard error for sample size n of σ/\sqrt{n} , where σ is the RRD standard deviation. Estimating the random errors for the sample standard deviation s , and for the Weibull and lognormal fit parameters is most easily done numerically. To do this, we generated 10 000 samples of 100 elements for each of our distributions. We then estimated the distribution parameters of the generating distribution using sample sizes of 5, 10, 20, 40, 80, and all 100 entries. We also fit the error (as measured by the standard deviation on the estimated fit parameter distribution) to a power law (Fig. 8).

Calculating the design-to Ibias assuming only random sampling errors and comparing to the results in Table I, we find that random sampling errors account for about 50%–70% of the increase in Ibias from best fit to 95% WC, assuming a Weibull fit. Similarly, we find that sampling errors account for about 60% of the increases assuming either a normal or lognormal fit.

Fitting the data for the 60 and 100 krad(Si) steps and adding it to the 200 krad(Si) data yields Fig. 9 for 99.9999% success probability at the 95% CL.

The linearity between dose and increased Ibias makes it possible to correlate the design-to Ibias with equivalent radiation design margins (RDM). The 99.9999/95 values are roughly $2.5 \times$ the mean values for each dose step. Factoring in the usual $2 \times$ margin, this means a $5 \times$ RDM ought to establish the same reliability level. While this procedure carries risk (e.g., due to process changes), it provides an empirical basis for the RDM, and is preferable to an across-the-board, ad-hoc requirement.

Like high reliability requirements, large flight lots emphasize the distribution tails. Failure in a flight lot of N parts with no redundancy is driven by the worst performing part. For failure pdf p_f (Ibias) and cumulative distribution P_f (Ibias), the pdf for the worst-case Ibias is

$$p(\text{Max Ibias}) = \binom{N}{1} p_f(\text{Ibias}) \times (1 - P_f(\text{Ibias}))^{N-1}. \quad (5)$$

Table II gives best fit and 95% WC design-to values for flight lots of 20 and 100 RH1014s.

The analyses carried out above demonstrate both the increased power and increased risk for of assuming a particular distribution form for the RRD. In contrast, an analysis with purely binomial statistics assumes all samples come from a common distribution, but makes no assumptions about the distribution form. For the RH1014 data considered here, such an analysis can conclude at the 95% CL only that fewer than 1.9% of the parts will exceed the highest measured Ibais for each dose level. The limitations of such an analysis become more apparent when we deal with smaller datasets.

B. OP-07 Op amp

For parts with smaller archival datasets, the procedure is the same, but the more limited statistics give rise to larger random errors. Our dataset for the Analog Devices OP-07 ultra-low offset voltage operational amplifier consists of measurements

$\mu \sigma$	7.7	7.9	8.1	8.3	8.5	8.7	8.9	9.1	9.3	9.5	9.7	9.9	10.1	10.3
55.2	-7.53	-6.19	-5.13	-4.33	-3.75	-3.37	-3.17	-3.12	-3.23	-3.46	-3.81	-4.26	-4.81	-5.44
55.4	-6.65	-5.35	-4.34	-3.57	-3.02	-2.68	-2.51	-2.49	-2.62	-2.88	-3.25	-3.73	-4.30	-4.95
55.6	-5.87	-4.62	-3.63	-2.90	-2.39	-2.07	-1.92	-1.94	-2.09	-2.37	-2.76	-3.26	-3.84	-4.51
55.8	-5.20	-3.98	-3.03	-2.32	-1.83	-1.54	-1.42	-1.45	-1.63	-1.93	-2.34	-2.85	-3.45	-4.13
56	-4.63	-3.43	-2.51	-1.83	-1.36	-1.09	-0.99	-1.04	-1.23	-1.55	-1.98	-2.50	-3.12	-3.81
56.2	-4.16	-2.99	-2.09	-1.42	-0.98	-0.72	-0.64	-0.70	-0.91	-1.24	-1.68	-2.22	-2.84	-3.55
56.4	-3.79	-2.64	-1.75	-1.11	-0.68	-0.43	-0.36	-0.44	-0.66	-1.00	-1.45	-1.99	-2.63	-3.34
56.6	-3.53	-2.39	-1.52	-0.88	-0.46	-0.23	-0.16	-0.25	-0.48	-0.82	-1.28	-1.83	-2.48	-3.20
56.8	-3.37	-2.24	-1.37	-0.74	-0.33	-0.10	-0.04	-0.14	-0.37	-0.72	-1.18	-1.74	-2.38	-3.11
57	-3.31	-2.18	-1.32	-0.69	-0.28	-0.06	0.00	-0.10	-0.33	-0.68	-1.14	-1.70	-2.35	-3.07
57.2	-3.36	-2.22	-1.36	-0.73	-0.32	-0.09	-0.03	-0.13	-0.36	-0.71	-1.17	-1.73	-2.37	-3.10
57.4	-3.50	-2.36	-1.49	-0.86	-0.44	-0.21	-0.14	-0.23	-0.46	-0.81	-1.26	-1.82	-2.46	-3.18
57.6	-3.75	-2.60	-1.72	-1.07	-0.64	-0.40	-0.33	-0.41	-0.63	-0.97	-1.42	-1.97	-2.61	-3.32
57.8	-4.11	-2.94	-2.04	-1.38	-0.93	-0.68	-0.60	-0.67	-0.87	-1.20	-1.64	-2.18	-2.81	-3.52
58	-4.56	-3.37	-2.45	-1.77	-1.31	-1.04	-0.94	-0.99	-1.19	-1.50	-1.93	-2.46	-3.08	-3.78
58.2	-5.12	-3.90	-2.95	-2.25	-1.77	-1.47	-1.36	-1.39	-1.57	-1.87	-2.29	-2.80	-3.40	-4.09
58.4	-5.78	-4.53	-3.55	-2.82	-2.31	-1.99	-1.85	-1.87	-2.03	-2.31	-2.70	-3.20	-3.79	-4.46
58.6	-6.54	-5.25	-4.24	-3.48	-2.94	-2.59	-2.43	-2.42	-2.55	-2.81	-3.19	-3.66	-4.24	-4.89
58.8	-7.41	-6.07	-5.02	-4.22	-3.65	-3.27	-3.08	-3.04	-3.15	-3.38	-3.73	-4.19	-4.74	-5.38

Fig. 7. ML Best fit (black) 90% (dark gray) and 95% (light gray) CL intervals for the normal distribution for data in Fig. 6. Numbers in each cell are the logarithms of the likelihood ratios relative to the maximum likelihood.

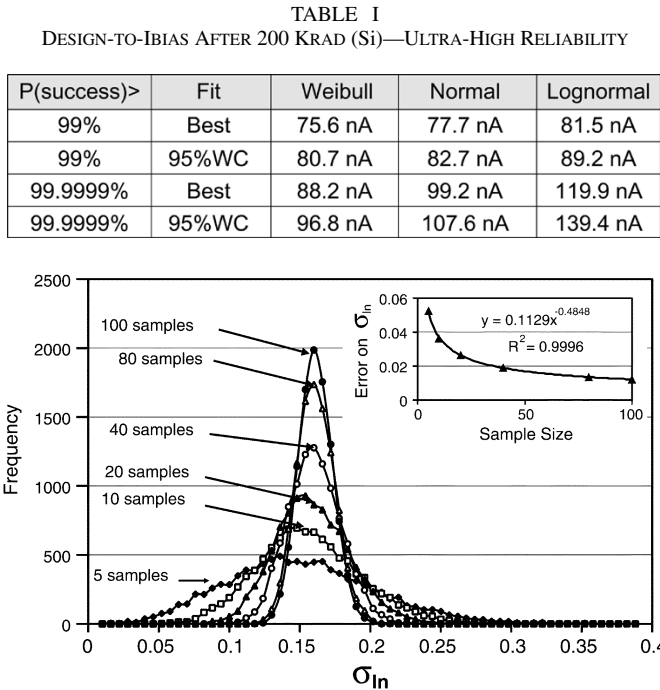


Fig. 8. Distribution widths of sample lognormal standard deviations σ_{In} , decrease with increasing sample size roughly according to power law (inset).

of excess input current (I_{in}) for 36 parts—traceable to two lot date codes and nine different wafers—after 100 and 300 krad(Si). The data for 100 krad(Si) are well behaved, indicating at the 95% CL that any pathology accounts for fewer than 8% of the parts. The Weibull, normal, and lognormal distributions all give reasonable fits to the data (Fig. 10).

Table III compares the best-fit and 95% WC fit parameters for the OP-07 and the RH1014, along with the percent changes between the best-fit and 95% WC parameters. Even though the OP-07’s best-fit distribution is slightly narrower than that for the RH1014, its Δ s in going from best fit to 95% WC are roughly twice those for the RH1014. Given that we have about $4 \times$ as much data for the RH1014 and that ML errors typically scale as

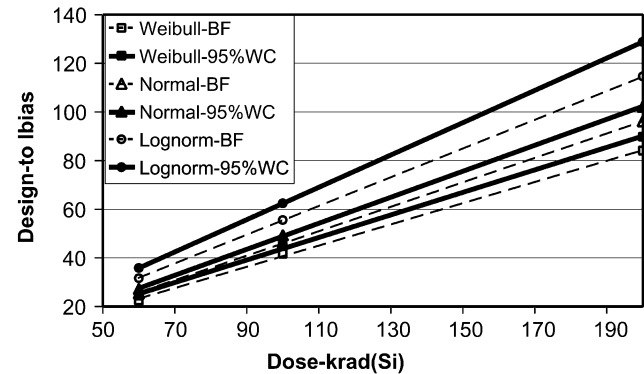


Fig. 9. Design-to Ibias for 99.9999% P(success)—best fit (open symbols) and 95% CL (solid symbols).

TABLE II
DESIGN-TO-IBIAS AFTER 200 KRAD (Si)—LARGE FLIGHT LOTS

P(s)/Fit	# of parts	Weibull	Normal	Lognormal
99%/Best	20	81.1 nA	86.2 nA	95 nA
99%/95%WC	20	87.7 nA	92.6 nA	106.5 nA
99%/Best	100	83.4 nA	90 nA	101.7 nA
99%/95%WC	100	90.6 nA	97 nA	115.2 nA

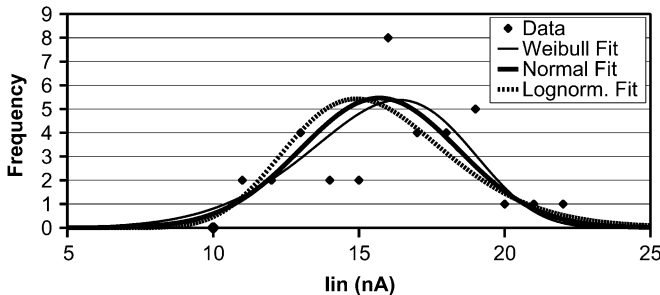


Fig. 10. Increased input leakage current for OP-07 op amps after 100 krad, and best fits for Weibull, normal, and lognormal forms.

the inverse square of the amount of data, this difference is about what would be expected. However, the practical consequences

TABLE III
ML FIT PARAMETERS: FROM BEST FIT TO 95% WC

	RH1014 BF	RH1014 95% WC	% Δ in going to 95% CL	OP-07 BF	OP-07 95% WC	% Δ in going to 95% CL
Weibull width	29.04	29.99	3.27%	16.8	18	7.14%
Weibull shape	6.88	5.88	-14.53%	6.5	4.5	-30.77%
Normal mean	27.2	28.15	3.49%	15.7	16.7	6.37%
Normal sigma	4.3	5.05	17.44%	2.7	3.8	40.74%
Lognormal mean	3.292	3.326	1.03%	2.74	2.8	2.38%
Lognormal sigma	0.161	0.189	17.39%	0.18	0.25	38.89%

TABLE IV
BEST-FIT AND 95% WC DESIGN-TO IIN (100 AND 300 KRAD(Si))

Design-to Iin (nA)	Weibull Best Fit	Weibull 95% WC	Normal Best Fit	Normal 95% WC	Lognormal Best Fit	Lognormal 95% WC
100 krad(Si), Ps>99%	21.3	25.1	22	25.5	23.5	29.3
300 krad(Si), Ps>99%	61.9	78.8	66	80.2	77.9	117.2
300 krad(Si), Ps>99%, w/outlier	88.5	128.5	82.5	104.9	90.2	141.9
100 krad(Si), Ps>99.9999%	25.2	31.9	28.6	34.6	36.3	53.5
300 krad(Si), Ps>99.9999%	77.8	110	91.7	116.2	157.4	289.8
300 krad(Si), Ps>99.9999% w/outlier	136.4	195.2	123.2	162	205.7	398.7

of this lower precision are that design to values will be more conservative.

At 300 krad(Si), the Δ Iin distribution shifts to the right and broadens so that $\mu \approx 41.25$ nA and $\sigma \approx 10.6$ nA, as long as one ignores a single “outlier” with Δ Iin = 121 nA. The presence of the outlier among 36 parts indicates that up to 12% (95% CL) of the parts may be part of an RRD pathology at this dose level. There is also the question of whether to omit the outlier from the analysis or consider it as part of the RRD. We calculate the design-to Iin values, first omitting and then including the putative outlier (Table IV).

We can draw several conclusions from Table IV.

- 1) Smaller datasets lead to higher design-to values.
- 2) The lognormal distribution produces worst-case values due to its positive skew.
- 3) Including the outlier leads to much higher design-to values and produces such a broad distribution that the Weibull shape $s < 3.68$ is skewed right rather than left, so the Weibull produces higher values than the normal.
- 4) If we exclude the outlier, the Weibull gives the best fit, while including that outlier means that only the lognormal gives a reasonable fit.

Finally, we consider the relation between design-to values and radiation design margin such as we employed for the RH1014. For the RH1014, the linear relationship between damage and dose facilitated such an analysis. The analysis is more problematic for the OP-07, since we have only two dose steps, making it impossible to confirm a linear relationship

between dose and damage. However, the moderate degradation at the 100-krad(Si) dose step suggests that a substitution of RDM for lot testing may be possible up to this level—with $3.8 \times$ margin equating to 95/99 assurance and $6.8 \times$ being required for 95/99.9999% assurance. Confidence in such a substitution could be increased if degradation below 100 krad(Si) was confirmed to be linear in dose. The outlier seen for the 300-krad(Si) dose step makes it unwise to substitute RDM for RLAT above 100 krad(Si).

VI. LARGE LOT-TO-LOT FLUCTUATIONS

When lot-to-lot variability greatly exceeds intralot variations, the ensemble exaggerates flight-lot variability. At the same time, sampling errors due to small RLAT samples preclude meaningful inference. The first step in such cases is to characterize the variability and use trends in the data (e.g., relations between distribution mean and width) to infer flight-lot performance. If no trends emerge, it may be useful to assume that the mean and width are independent, estimating width with archival data and inferring flight-lot mean with lot-specific data. There are several advantages to this procedure. First, it is physically plausible, since shifts in the mean radiation response often result from lot-specific process changes, while intra-lot variations may be due to process tolerances. Second, while the sample mean for small samples converges rapidly to the parent-distribution mean, convergence of the variance of the distribution of sample variances (s^2) has a more complicated dependence on sample size n

$$\text{var}(s^2) = (1/n) \left[M_4 - \frac{n-3}{n-1} \sigma^4 \right] \quad (6)$$

where M_4 is the fourth central moment and σ the standard deviation of parent distribution. As n increases, (6) approaches convergence as n^{-1} . For a Normal RRD, sample sizes required to determine the sample mean within an error of σ are four parts for 90% confidence, five parts for 95% confidence, and nine parts for 99% confidence. Required samples scale as the inverse square of the allowable error. For other distributions, the sample size can be estimated numerically, but the same rule applies: parameters that most affect where the distribution is centered (e.g., normal mean, lognormal mean, and Weibull width) will generally converge rapidly even for small sample sizes, while those that affect mainly the distribution width and shape (e.g., normal and lognormal σ or Weibull shape) will benefit most from the additional statistics supplied by the ensemble distribution.

Mean shifts in gain Δ hFE for 2N2907 PNP transistors after 300 krad(Si) vary significantly from lot to lot, but are grouped tightly within a lot (especially for lots 1–4, 9, and 11, for which all parts are from a single wafer). Fig. 11 plots the inverse of the shifts for each lot. We fit each lot to a normal distribution using an ML fit in the usual manner—determining the mean and standard deviation σ independently for each lot. We also fit the data allowing the mean to vary from lot to lot, but assuming a common value for σ across all lots. In Table V, we see that the MLE for σ (column 2) varied considerably from sample to sample. If we take the 95% WC estimate, the variation is even greater. On the other hand, comparing the MLE for the common

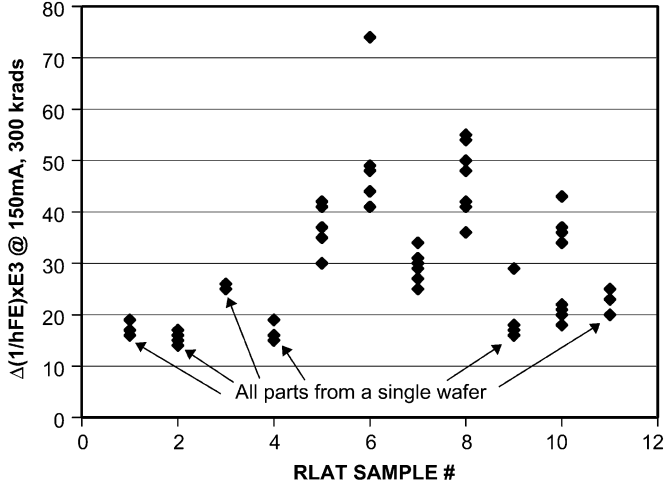


Fig. 11. Gain shifts in 2N2907 PNP transistors are consistent within wafer lots, but vary significantly across lots. Parts in lots 1–4 and 9 are from a single wafer.

TABLE V
COMPARISON OF ENSEMBLE AND SINGLE-LOT σ

Type of Lot	Best-Fit σ Range	95% CL WC σ Range	Best-Fit σ Across Lots	95% CL WC σ Across Lots
Single-Wafer	0.15–5.3	1.7–11.8	2.1	3.6
Multi-Wafer	2.7–11.8	6–23.3	6.8	9.8

σ across lots (column 4) and the 95% WC common σ , we see there is significantly less variation. As the required confidence level increases, the reduced random sampling errors due to the greater statistics of the ensemble provide a clear advantage.

VII. EFFECTS OF REDUNDANCY

Effectively implemented redundancy can significantly increase reliability not just for random failures, but in some cases even for wear-out type failure mechanisms and TID degradation. If instead of requiring all N parts in the flight lot to pass some criterion (e.g., $I_{bias} < \text{design} - \text{to value}$), only M parts must pass (N for M redundancy), failure will be driven by the $(M + 1)$ th worst part, and the failure pdf will be given by

$$p(I_{bias_{M+1}}) = \binom{N-M}{1} \binom{N}{N-M} (P_f(I_{bias}))^M \times p_f(I_{bias}) \times (1 - P_f(I_{bias}))^{N-M-1}. \quad (7)$$

By ensuring that the failure rate is no longer driven by the worst parts, added redundancy de-emphasizes the importance of distribution extremes. As Fig. 12 shows, when the system approaches 100:50 redundancy, the failure distribution of the redundant system is driven by the median behavior of the single-part distribution—even if the single-part distribution is bimodal or thick-tailed. A small test sample is much more likely to elucidate a distribution's central behavior than that of its tails. If an application requires large flight lots or high reliability, but testing large samples is impractical, redundancy may be the best way to increase confidence in mission success—and to simplify qualification.

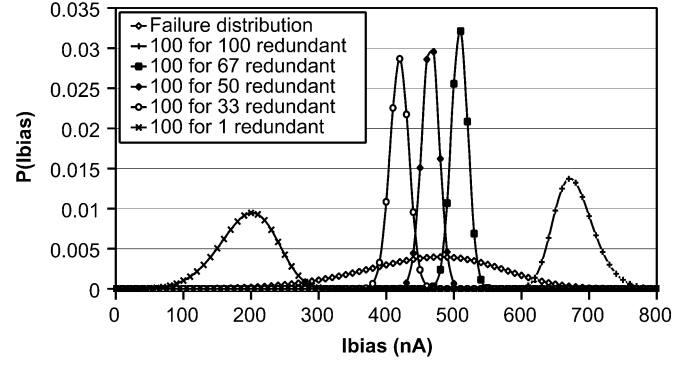


Fig. 12. For a hundred-part flight lot with a Weibull single-part failure distribution ($w = 500$, $s = 5.3$), increased redundancy causes the system failure distribution to approach and then surpass the median behavior of the single-part failure distribution (median behavior is achieved at 100:50 redundancy).

While redundancy can greatly improve system reliability and easing qualification, the efficacy of the redundancy depends on how it is implemented. First, the analysis above uses binomial statistics, and so it assumes that all of the parts have a common RRD. This may not be the case if some parts in the system have different application conditions (e.g., parts with bias-dependent radiation response). Also, the example illustrated in Fig. 12 assumes true 100:50 redundancy—that is, any 50 parts may fail without causing system failure. Had the system been implemented instead with 50 2:1 redundant pairs (failure of both parts in any pair causes a system failure), system reliability would still have been increased, but the system failure distribution would not have been centered on the single-part median, making it more difficult to characterize system reliability with a small sample test.

VIII. CONCLUSION

Large flight lots and high reliability requirements pose problems for RHA methods that rely on small sample sizes. Sampling errors arising from the small sample size and systematic errors arising from the assumed form of the RRD can invalidate RHA analyses. In this work, we have shown that sampling errors can be bounded by augmenting RLAT results with archival data to characterize radiation response variability and reduce random errors. Systematic errors from the assumed distribution form can be estimated by fitting the data to multiple forms—as we have done here for the normal, lognormal, and Weibull distributions. The results for the different distributions gauge the sensitivity of the analysis to the assumed distribution form. If there is no reason to favor one distribution over the others, it is prudent to assume the distribution that yields worst-case results.

Using archival data to bound distribution pathologies and infer performance is simplest when inter-lot and intra-lot variability are comparable. Then, the greater statistics of the ensemble distribution allow inference of flight-lot performance with greater confidence than would be possible with RLAT data alone. When lot-to-lot variability dominates, the first step is to look for trends in the data that are useful for prediction. Otherwise, we use RLAT data to estimate the mean (or other parameter

most affecting distribution central behavior) and the ensemble to estimate the standard deviation (or other parameter most affecting distribution shape and width).

In looking at the effects of large flight lots, it is important to use a distribution that considers flight lot size (5) and system redundancy if present (7).

We have also noted that redundancy not only may increase system robustness to random failures, but also to degradation and wear-out-type failure mechanisms. Redundancy can simplify part qualification by driving system performance toward the central portion of the radiation response distribution—a region much more easily characterized by small sample sizes. Early in a part's history, when archival data are not available, the only options for use in high reliability or large-flight-lot applications may be qualification using very large samples or conservative design including significant (2n:n or more) redundancy.

REFERENCES

- [1] R. Pease, Microelectronic piece part radiation hardness assurance for space systems, Atlanta, GA, July, 19. 2004 NSREC Short Course.
- [2] *Ionizing dose and neutron hardness assurance guidelines for microcircuits and semiconductor devices*, 2005. MIL-HDBK-814.
- [3] A. Johnston and C. Lancaster, "A total dose homogeneity study of the 108 A operational amplifier," *IEEE Trans. Nucl. Sci.*, vol. NS-26, no. 6, pp. 4769–4774, Dec. 1979.
- [4] J. Krieg *et al.*, "Hardness assurance implications of bimodal total dose response in a bipolar linear voltage comparator," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 6, pp. 1627–1632, Dec. 1999.
- [5] M. Shaneyfelt *et al.*, "Impact of passivation layers on enhanced low-dose-rate sensitivity and pre-irradiation elevated-temperature stress effects in bipolar linear ICs," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 6, pp. 3171–3179, Dec. 2002.
- [6] A. Namenson, "Lot uniformity and small sample sizes in hardness assurance," *IEEE Trans. Nucl. Sci.*, vol. 35, no. 6, pp. 1506–1511, Dec. 1988.
- [7] (2005) Resources for Distributions, Likelihood Methods and Inference Include NIST/SEMATECH e-Handbook of Statistical Methods. [Online]. Available: <http://www.itl.nist.gov/div898/handbook/mathworld.wolfram.com, www.Weibull.com, www.resacorp.com>